Procedia Computer Science

Volume 88, 2016, Pages 277–281

7th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2016

Particular Qualities of the Semantic Web Training Course

Anita Balandina, Anastasiya Kostkina, Artyom Chernyshov, Boris Shchukin and Valentin Klimov

National research Nuclear University "MEPhI", Moscow, Russian Federation

anita.balandina@gmail.com, anakost@bk.ru,

a-chernyshov@protonmail.com, tsh-k22@mail.ru, vvklimov@mephi.ru

Abstract

In order to ease the lives of authors, editors, and trees, we present a manual and an example of the use of Microsoft Word and similar tools for creating documents for EasyChair-based electronic and on-paper publishing for the Procedia publication series by Elsevier Science.

Keywords: Semantic web, web-services, educational course.

1 Introduction

Over the past decade, in the information technology area mankind has achieved considerable success. Creation of the World Wide Web in 1993 by British scientist Tim Berners-Lee marked the beginning of a new milestone in the history of computer science and has served as a powerful incentive to speed up the integration of various processes, including the process of information globalization [1]. Therefore, there is a need for technologies that could implement the exchange of data with sufficient speed at various points of the globe. As a result, it was developed the set of data transfer protocols, languages, integration of architectural concepts. At the same time there is a development of many programming languages, such as Java, C #, now frequently used to develop client applications and server software.

Further, due to the developed technologies, the interaction between the components of an information system became possible, it means, that services located in different network nodes, often remote each other over thousands of miles, was able to interact between one another. There were the first distributed applications whose components can interact at a distance, using the protocol RPC. In this sense, the concept has gained popularity of service-oriented architecture(SOA), which is still used for the development of distributed applications.

Improving older technologies has led to the emergence of new ones, and therefore in the late 90's - early 2000's there was a technology called web-service. Web-service is a network technology that

Selection and peer-review under responsibility of the Scientific Programme Committee of BICA 2016

doi:10.1016/j.procs.2016.07.435

provides machine-to-machine cooperation based on web standards. A similar definition was formulated by a consortium W3C, involved in the development of web standards, according to which the Web service is "a software system designed to support interoperable machine-to-machine interaction over a network".

For the moment the technology of web services is well developed and widely used, providing a solution to a number of problems which the majority of software developers faced to. However, its closed standards and formats normally create problems in application integration. Web services are widely used, as they are based on open communication protocols (e.g., SOAP, HTTP) and transmission standards, and supported on various software platforms (WSDL, UDDI).

Implementation of the concept of the Semantic Web as an extension for World Wide Web (WWW) began in the early 2000s, but the process stretched out for many years, taking a new turn only in the last five years. Semantic Web provides an opportunity to interpret the information, as a person and a computer, which greatly improves the search for information.

Web services play a very significant role in the development of the Semantic Web. In this concept all web services should be provided with clearly described semantics, accessible via the Inter-net, and suitable for automated search, composition and execution. These web services were called Semantic Web Services. The interpretation the semantics of data is possible due to important characteristics - the use of the uniform resource identifier (URI), and the use of ontologies and metadata description languages. For the Semantic Web technology this stack includes a set of standard description languages, including XML, XML Schema, RDF, RDF Schema, OWL and others.

According to the carried-out analysis of state-of-the-art in the area of the Semantic Web, there are some certain problems impeding development of idea. Tim Berners-Lee, the author of the Se-mantic Web idea, names the existing methods of the organization of information in Web as “a simple idea, which is still largely unrealized” [2]. He also says that today there are not the public and available ways for viewing and direct usage of the information provided by sites of the Semantic Web.

2 Problems of the Semantic Web

In spite of all the advantages of the Semantic Web, this concept still has plenty of problems, which are mostly caused by underdeveloped technologies and standards. The most important problems are listed below [3].

1. Availability of semantic content is the important problem on the way to formation and usage of knowledge spaces, because the most of the information in the Web is not represented in the Semantic Web-formats and there's no hope that this work can be performed manually.
2. Many experts suppose that ontologies are the key component in solution of the problem of semantization of Web content. In this regard problems of onto-engineering (methods and development tools and evolutions of ontologies) and availability of already existing ontologies are of special importance. The significant efforts have to be made for storage, processing and search of semantic content, and the solutions in this area have to ensure the effective work with huge volumes of knowledge.
3. The problem of the multilingual content exists in the classic Web, but it also is one of the main problems for the Semantic Web, because the Semantic Web must support the access to the information is no matter which language it was originally presented.
4. The presentation of information for users (visualization of content) also has to undergo the essential changes and to provide free orientation in the huge number of the facts which meet his requirements.

5. The Semantic Web could be confronted by concerns regarding censorship and privacy. For example, text-analyzing techniques can now be easily bypassed by using other words, metaphors for instance, or by using images in place of words. An advanced implementation of the semantic web would be the big opportunity most of all for the government, to control the viewing and creation of online information, as this information would be much easier for an automated content-blocking machine to understand. Besides, the issue has also been raised that, with the use of FOAF files (vocabularies) and geolocation meta-data, there would be very little anonymity associated with the authorship of articles on things such as a personal blog.
6. The need to describe the metadata somehow leads to duplication of information. It would be much more time-consuming for creating and publish content because there would need to be two formats for one piece of data: one for human viewing and one for machines. However, many web applications in development are addressing this issue by creating a machine-readable format upon the publishing of data or the request of a machine for such data. The development of micro formats has been one reaction to this kind of criticism [4].
7. The last problem is connected with ensuring stability of the Semantic Web, and it, in turn, assumes that serious efforts have to be made in the area of the standardization which must provide creation of the technologies necessary for formation of knowledge spaces.

These problems raise doubts about the utility of the Semantic Web. Some critics point out that the development of the Semantic Web is too slow due to the fact that the followers of the ideas in particular W3C Consortium are not keeping pace with the times and they don't take into account the needs of developers and users. The other point of view is that there are some statements [4,5] of that the Semantic Web is developing quite rapidly and its influence has increased considerably over the last decade.

The current research given by PewInternet, a project of the Pew Research Center [6], represents the both points of view: approximately 40% of responses agree that by 2020, the Semantic Web envisioned by Tim Berners-Lee will not be as fully effective as its creators hoped and average users will not have noticed much of a difference. And, vice versa, about 50% of responses disagree with the above-mentioned statement and guess that by 2020, the Semantic Web envisioned by Tim Berners-Lee and his allies will have been achieved to a significant degree and have clearly made difference to the average Internet users. Other 10% didn’t respond.

The problems listed above could be easily solved by the qualified experts in the Semantic Web area. And here is another problem: the lack of the appropriate specialists. The majority of developers doesn’t know or simply understand the advantages of the Semantic Web concept. Therefore, we need to interest young professionals in the described concept, to make them continue the development and standardization of the Semantic Web.

In this connection it was decided to create an educational course, which would be aimed at the students, studying in the NRNU MEPhI on the Cybernetics Department.

3 Description of the Course

The developing course should contain all the topics connected with Semantic Web technologies, as well as web services technologies. At the moment there are two approaches to developing web services, which are mostly in use.

The first method is connected to SOAP (Simple Object Access Protocol). It is a protocol specification, which is used for the structured information exchange in the implementation of web services. It uses XM for its message format, and relies on other application layer protocols, such as HTTP, TCP/IP, etc.

The other one is REST (representational state transfer). It is the software architectural style of the World Wide Web. [7] REST gives a coordinated set of constraints to design the components of a distributed system.

Therefore, it was decided to include these two technologies to the concept of the course.

The course is divided into two parts: web services and Semantic Web.

The first part is connected to general web-services standards and technologies. Is consists of three laboratory works. The first one is devoted to the basics of the extensible markup language XML and language for describing the structure of XML-documents – XML-Schema (XSD). The goal of this work is to make the students familiar with this standards and syntax to make them able to use this knowledge on practice and during next laboratory works. The task consists of creating own XML-document describing the exact domain knowledge (their descriptions students will find in appendix), checking it via browser and generating the appropriate schema. The work has two variants. The first one is to complete the task using the MS Visual Studio XML editor, and the second one is to use XMLPad, an open-source XML editor.

The second and third works are aimed to introduce the students the most popular standards of web-services development. The second work is devoted to classical SOAP web-services. Students will have to create their own service according to their variant and domain knowledge. The first variant of the work is to develop WCF service using MS Visual Studio IDE and MS SQL Server, and the other one is to develop JAX-WS service using NetBeans IDE and MySQL. The students will have to explain the teacher principles of work of the developed service and its WSDL-description.

The third work is devoted to RESTful services. Students will have to develop web-service following the principles of REST architecture. The variants of work are almost the similar with the previous one: the first variant is to create ASP.NET Web API service; the second one is to create Java Web service. The goal of these two works is not only to introduce students these approaches, but show the difference between them and advantages and disadvantages of each method.

The second part of course is devoted to Semantic Web standards and technologies. It consists of three laboratory works.

The fourth laboratory work is connected to RDF and SPARQL. The students will have to create their own RDF-document describing their domain of knowledge. As an editor for RDF students are offered to use dotNetRDF, a freeware toolkit for work with RDF. Students will have to start creating their document in RDF/XML notation. After that, they will have to convert this description to other formats: N3, N-Triple and Turtle, and explain the differences between these notations to the teacher. Students also will have to visualize their RDF data by using online services, for example the one provided by W3C. This will help them to understand the RDF-document structure. The last task is to write a few SPARQL-queries to the described data. We suppose, that this laboratory work will fully cover the important Semantic Web standards related to RDF and make students familiar with them.

The fifth work is devoted to OWL. The students will have to create their own ontology according to their domain of knowledge. The recommended system for doing this work is Protégé editor. After completing the first task, students will have to check and export their ontologies via Protégé and explain in to the teacher. It's important to make students understand the process of creation the ontologies rather than just copy and paste the code.

Finally, the last work is devoted to OWL-S. The work consists of importing WSDL description of the service, which was developed during the second work, creating and checking web-service ontology. The recommended software is OwlWiz, a freeware OWL-S editor. After competing these tasks, the students will have to explain the teacher their ontologies.

Besides the tasks, each work contains the teaching materials, schemas, recommendations and examples of realizations these works, which include detailed information and screenshots. It also includes the control questions after each work, which could be used for knowledge control. The described laboratory work course will be accompanied with lections. At the moment, the practice book is in the process of publication, and there are great plans about its improvement and extension.

4 Conclusions

During the current work we developed an educational course, which consists of six laboratory works, in order to resolve the problem of student education for the Semantic Web technologies. Applying this course will enhance the education quality of department, level of graduating students and in perspective would solve some of the Semantic Web problems described above.

Currently, practice book "Web-services and Semantic Web" for our educational course is in the stage of publication.

We also developed a system, that can be useful in conducting out these labs. It currently supports only fourth laboratory work and provide support for RDF and SPARQL requests edit and RDF graph visualization.

As further work, we're planning to add an extra laboratory work – "Composition of Semantic web-services". This work would be devoted to different types of web-services composition [8].

Also, we will provide support for fifth work "OWL" and six work "OWL-S" in our system and will expand it with API that can be used to integrate our system with any distance learning system, including the Cybernetics department system.

In order to show the effectiveness of the developed course we will enroll it on the students studying in the NRNU MEPhI on the Cybernetics Department.

Acknowledgements

This work was supported by Competitiveness Growth Program of the Federal Autonomous Educational Institution of Higher Professional Education National Research Nuclear University MEPhI (Moscow Engineering Physics Institute). The funding for this research was provided by the Russian Science Foundation, Grant RSF 15-11-30014.

References

[1] Markov A.A., Molchanova O.I., Polyakova N.V. *Mass information theory and practice*. (2014). *Moscow: Infra-M*, 251 pages.

[2] Shadbolt N., Berners-Lee T., Hall W. The Semantic Web Revisited. (2006). *IEEE Intelligent Systems*, 3, 96-101.

[3] Tzitzikas Y.,Kotzinos D. (Semantic Web) evolution through change logs: Problems and solutions. (2007). *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications*, 610-615.

[4] Livesey A.J. The Semantic Web as a Knowledge Management Environment. (2013). *Stellenbosch University. Faculty of Arts and Social Sciences. Dept. of Information Science*, 110.

[5] Dimitrov M. Crossing the Chasm with Semantic Technologies. (2014). *WaSABi-FEOSW*, 1240, 166-167.

[6] Anderson J., Rainie L. The Fate of the Semantic Web. Prospects for the semantic web. (2010). *PewInternet, Available at: http://www.pewinternet.org/2010/05/04/prospects-for-the-semantic-web/*

[7] Richardson L., Amundsen M. RESTful web APIs. (2013). *O'Reilly Media*, 404 pages.

[8] Klimov V. V., Chernyshov A. A. Shchukin B. A. Composition of Web-services Based on Semantic Description. (2015). *Proceedings of the 11th International Conference on Web Information Systems and Technologies Lisbon*, 124-128